Investigating D-Lysine Stereochemistry for Epigenetic Methylation, Demethylation and Recognition

Roman Belle,[a,b‡] Abbas H. K. Al Temimi,[a‡] Kiran Kumar,[b] Bas J. G. E. Pieters,[a] Anthony Tumber,[b] James E. Dunford,[c] Catrine Johansson,[bc] Udo Oppermann,[c] Tom Brown,[b] Christopher J. Schofield,[b] Richard J. Hopkinson,[b] Robert S. Paton,[b] Akane Kawamura,[b] and Jasmin Mecinović*[a]

[a] Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands. E-mail: j.mecinovic@science.ru.nl; Fax: +31 24 3653393; Tel: +31 24 3652381

[b] Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, OX1 3TA Oxford, UK

[c] Botnar Research Centre, NIHR Oxford Biomedical Research Unit, University of Oxford, Windmill Road, OX3 7LD Oxford, UK

† Electronic supplementary information (ESI) available: Supporting figures, enzyme assays, peptide synthesis, protein production, synthesis, NMR data.

‡ These authors contributed equally to this work.

Histone lysine methylation is regulated by N$^{\varepsilon}$-methyltransferases, demethylases, and N$^{\varepsilon}$-methyl lysine binding proteins. Thermodynamic, catalytic and computational studies were carried out to investigate the interaction of three epigenetic protein classes with synthetic histone substrates containing L- and D-lysine residues. The results reveal that out of the three classes, N$^{\varepsilon}$-methyl lysine binding proteins are superior in accepting lysines with the D-configuration.

Histone tails are subject to a plethora of posttranslational modifications (PTMs); acetylation, methylation, phosphorylation[1] are established with many more recently discovered.[2–4] Histone lysine methylation is linked to both gene activation and repression, depending on the methylation state and modification site. Methylation is catalysed by S-adenosylmethionine (SAM) dependent histone lysine methyltransferases (HKMT) that install one (Kme1), two (Kme2), or three (Kme3) methyl groups on the lysine N$^{\varepsilon}$-amino group (Fig. 1).[5] Nε-Methyl group removal is catalysed by flavin-dependent lysine specific demethylases (KDM1 in humans) that accept only Kme1 or Kme2 modifications, or by the Fe(II) and 2-oxoglutarate (2OG) dependent JumonjiC (JmjC) demethylases (KDM2-7) that catalyse demethylation of all three types of Nε-methylated lysines (Fig. 1).[6] Methylated lysines are recognised by different classes of reader proteins, including plant

homeodomain (PHD) zinc fingers, tandem tudor domains (TTD), chromodomains (CD) and malignant brain tumor (MBT) domains (Fig. 1).[7] In general, hydrogen bonds and electrostatic interactions appear relatively more important in binding of Kme1 and Kme2 to the reader proteins, whereas Kme3 groups bind to aromatic cage-containing reader proteins via cation–π interactions[8,9] involving displacement of water molecules.[9]

Structural and mechanistic work implies that three classes of histone N^{ε}-methyl lysine interacting proteins accept the N^{ε}-(methyl) L-lysine as their natural substrate/ligand; however, their selectivity with respect to lysine C^{α} stereochemistry has not been investigated. This is of interest because protein residue epimerisations and methylation patterns are implicated in ageing[10]/disease[6] and because some 2OG oxygenases, including some JmjC KDMs and hydroxylases, have a broad substrate selectivity.[11] Along with rearrangements, including Asn and Gln residues, D-amino residues have been observed in ageing/diseased cells[12] and it has been proposed that D-lysine residues occur in tumour cells.[13]

We hypothesised that there might be a selectivity between the three classes of epigenetic proteins for the acceptance of (methylated) D-lysine residues on histones. Here, we report studies comparing the selectivity of histone lysine methyltransferases, histone lysine demethylases and epigenetic readers for (methylated) L- and D-lysine residues.

We first used isothermal titration calorimetry (ITC) to investigate the binding of representative PHD zinc fingers and TTDs recognising the H3K4me3 mark, i.e: i) the PHD zinc fingers of $KDM5A_{PHD3}$, $TAF3_{PHD}$, and $BPTF_{PHD}$ and ii) the TTDs of $SGF29_{TTD}$ and $KDM4A_{TTD}$.[7] Thermodynamic parameters for the associations between reader proteins and 10-mer histone peptides that bear L-Kme3 and D-Kme3 were measured. In all cases, the readers form stronger complexes with the L-Kme3, compared to the D-Kme3-histone sequences (Table 1, Fig. S1 in ESI). D-H3K4me3 bound to $KDM5A_{PHD3}$, $TAF3_{PHD}$ and $BPTF_{PHD}$ with an ~10-fold lower affinity than L-H3K4me3, whereas the association of D-H3K4me3 to $SGF29_{TTD}$ and $KDM4A_{TTD}$ decreased more substantially (~30–35 fold, compared to L-H3K4me3). The results reveal that decreased affinity for D-H3K4me3 relative to L-H3K4me3 derives from less favourable enthalpy of binding (ΔH°); values of $\Delta\Delta H^{\circ}$ are in the range of 2.1 kcal mol-1 (for $TAF3_{PHD}$) to 10.5 kcal mol-1 (for $KDM4A_{TTD}$) (Table 1). The entropy of binding ($-T\Delta S^{\circ}$) is relatively more favourable for D-Kme3-containing histone peptides, but this does not compensate for the decreased enthalpy values (Table 1). We have reported that binding of 10-mer H3G4 (where L-Kme3 is substituted by glycine) results in a large decrease (>500-fold) in binding strength to the same readers.9 The observation that replacement of L-Kme3 by D-Kme3 causes a smaller reduction in affinity (8–36-fold) and enthalpy suggests that the D-Kme3 side chain is involved in energetically favourable interactions, possibly within the aromatic cage; it is also possible that binding of D-Kme3 leads to less favourable interactions of neighbouring amino acids in the histone peptide with the reader proteins.

Temporal atomistic molecular dynamic (MD) simulations were then used to investigate the relative flexibility of the readers and their ability to accommodate D-Kme3. MD approaches have proven to be valuable for elucidating properties concerning binding of histone PTM residues by epigenetic proteins.[14,15] Previously, MD simulations (10 ns with the Amberff99sb force field) of

H2AK5ac, H4K12ac, and H3K14ac binding to the BRPF1 bromodomain were used to understand selectivity.[14] Binding modes to the multi-domain JARID demethylase protein, KDM5C, have also been studied by MD (10 ns with Amberff14sb).[15]

We simulated epigenetic reader proteins with the L-Kme3 and modified D-Kme3 residues. Starting structures were taken from representative crystal structures where the L-Kme3 stereocenter was manually inverted to generate the D-Kme3 complex, with priority given to replicating the position of L-Kme3 in the original PDB structure (see ESI). All simulations used AMBER12.[16] Systems were solvated in a 10 Å truncated octahedral box of TIP3P[17] water, neutralised explicitly with either sodium or chloride ions, and simulated for a total of 10 ns. The binding poses of the L- and D-Kme3 residues relative to the surrounding reader aromatic cage following minimization and equilibration are shown in Fig. S2. Qualitative analysis of snapshots taken after 5 ns and 10 ns of MD simulation show the behaviour of the complexes over time (Fig. S3-7). Orientations of the two Kme3 stereoisomers within the aromatic pocket of the same reader are similar at 0 ns (Fig. S2A-D), with the exception of $SGF29_{TTD}$ (Fig. S2E). $KDM5A_{PHD3}$ bound to both D- and L-Kme3 exhibit the most similar pose regarding placement of the W18-W28 aromatic cage and modified residues (Fig. S2A). In the case of $SGF29_{TTD}$, D-Kme3 more fully occupies the aromatic Y238-Y245-F264 cage than does L-Kme3; however, L-Kme3 adopts a more similar orientation to D-Kme3 by 5 ns that is observed throughout the simulation (Fig. 2SE). For the two H3-PHD reader complexes, $KDM5A_{PHD3}$ (Fig. S3) and $TAF3_{PHD}$ (Fig. S4), major differences in the H3 chain backbone geometry were observed throughout the simulation when comparing the L- and D- systems (Table S1). The reorientation of H3 backbone for D-H3K4me3 bound to $TAF3_{PHD}$ and $KDM5A_{PHD3}$ is stabilized by favourable cation-π interactions with the W868-W891 and W18-W28 cages, respectively. Movement of the H3 backbone is minimal for the other 3 reader complexes (Fig. S5-7). We analysed the electrostatic energies (E_{ele}) between N^{ε} of L-Kme3 and D-Kme3 and the π-system of reader protein aromatic cage residues during each simulation. Electrostatically-dominated cation-π interactions are present for both stereoisomers. Those involving L-Kme3 are slightly more favourable than D-Kme3 for all reader proteins except $SGF29_{TTD}$ (Table S2). This can be explained by the non-optimal starting geometry (Fig. S2E), wherein L-Kme3 is pointing away from the Y238-Y245-F264 aromatic cage of $SGF29_{TTD}$.

Overall, the combined MD simulations and ITC binding studies indicate the ability of reader proteins to efficiently accommodate D-H3K4me3 residues. Electrostatic energy calculations between N^{ε} of Kme3 and surrounding aromatic residues suggest the presence of favourable stabilizing cation-π interactions for both stereoisomers. Examination of the H3 orientation for D-Kme3 complexed with readers $KDM5A_{PHD3}$ and $TAF3_{PHD}$ emphasize the flexibility of the histone backbone to potentially prioritize the cation-π interactions. The observed reduction in binding enthalpy for D-Kme3 is presumably due to weaker cation–π interactions, weaker H-bonding, and/or the release of smaller number of water molecules in the aromatic cage. The more favourable entropy of binding for D-Kme3 may arise from a higher conformational degree of freedom of D-Kme3 side chain (than L-Kme3 side chain) in the complex.

We then explored whether histone lysine methyltransferases catalyse methylation of D-lysine residues. Four representative human methyltransferases were chosen to investigate D/L-

stereoselectivity: SETD7 that monomethylates H3K4, SETD8 that monomethylates H4K20, and G9a and GLP that di- and trimethylate H3K9.[5] MALDI-TOF MS analyses of SETD7-catalysed methylation of H3K4 manifested quantitative production of monomethylated H3K4 (i.e. L-Lys at position 4) under standard assay conditions. In contrast, D-H3K4 (i.e. D-Lys at position 4) was not observed to be a substrate under the same conditions (2 μM SETD7, 100 μM histone peptide, 200 μM SAM, pH 8.0, 37 °C, 1 hour) (Fig. 2A, S8-9). Increased amounts of SETD7 (10 μM) and SAM (1 mM), and prolonged incubation (3 hours) did not yield any monomethylated product D-H3K4me1. Similarly, SETD8 was observed to efficiently catalyse monomethylation of H4K20, but no methylation of D-H4K20 was detected under standard conditions (Fig. 2B), or with prolonged incubation with additional SETD8/SAM (Fig. S10-11). MALDI-TOF MS assay of G9a/GLP-catalysed methylation of 15-mer H3K9 peptide revealed efficient formation of H3K9me3 (Figs. 2C,2D and S12,13). Under the same conditions, we only observed trace evidence for monomethylation of D-H3K9 with G9a/GLP. A significantly larger amount (20-35%) of the D-H3K9me1 product was observed with increased G9a/GLP (10 μM) and SAM (1 mM) after 1 hour (37 °C); all 3 methylated products were observed after 6 hours (Fig. S14-15). Competitive experiments between 14-mer L-H3K9 and 15-mer D-H3K9 indicated that D-H3K9 does not inhibit G9a-catalysed methylation of H3K9 within detection limits (Fig. S16). Collectively, the results imply the tested methyltransferases are specific for the L- over the D-lysine stereoisomers. Notably, D-lysine is poorly accepted by some methyltransferases (G9a and GLP). Structural analyses of histone lysine methyltransferases with substrates reveal binding in a narrow apolar tunnel;[18] it is likely that the positioning of D-lysine in this tunnel is non-optimal, thus slowing catalysis (Fig. 1B).

We investigated potential demethylation of methylated D-lysine residues by histone lysine demethylases, using 6 human histone lysine demethylases, i.e. catalytic domains of KDM1A (that demethylate H3K4me2), $KDM5B_{JmjC}$ and $KDM5C_{JmjC}$ (that demethylate H3K4me3), and $KDM4A_{JmjC}$, $KDM4D_{JmjC}$ and $KDM4E_{JmjC}$ (that demethylate H3K9me3).[6] LC-MS analyses confirmed activity with the L-lysine H3K4me2/H3K4me3/H3K9me3 peptides. Thus, under standard conditions (200 nM enzyme, 5 or 10 μM substrate), we observed KDM1A-catalysed demethylation of H3K4me2 to H3K4, $KDM5B_{JmjC}$/$KDM5C_{JmjC}$-catalysed demethylation of H3K4me3 to H3K4me2 and H3K4me1, and $KDM4A_{JmjC}$/$KDM4D_{JmjC}$/$KDM4E_{JmjC}$-catalysed demethylation of H3K9me3 to H3K9me2 and H3K9me1 (Fig. 3, S17-23). Under the same conditions, the results of analogous D-lysine substrates showed only traces (at most) of possible demethylation (Fig. 3, S17-23), implying that histone lysine demethylases require the L-stereochemistry for efficient catalysis. Use of higher enzyme concentrations (600 nM or 1 μM) did not enhance demethylation with methylated D-lysine, yielding only traces of demethylated products.

To test for potential stimulation of 2OG turnover by D-H3K9me3 we employed 1H NMR, monitoring signals of H3K9me1/2/3, 2OG and succinate catalysed by $KDM4E_{JmjC}$. The results confirmed that $KDM4E_{JmjC}$ catalyses L-K9me3 demethylation (Fig. S24,26).[19] Analysis of the D-K9me3 potential substrate shows substantial lower formation of Kme1/2 under the same conditions, compared to L-K9me3 (Fig. S25,27). Use of an increased enzyme did not increase formation of D-Kme1/2 for D-K9me3. (Note, traces (≤5%, determined by 1H NMR (Fig. S27)) of

demethylation for the D-Kme3 may be due to the presence of very low levels of the L-epimers produced during synthesis.) Initial rates of uncoupled 2OG turnover (Fig. S26-28) do not indicate substantial differences between D-K9me3 and L-K9me3, suggesting that D-K9me3 does not stimulate significant 2OG turnover. Additional experiments to examine the influence of the D-K9me3 on L-K9me3 demethylation in a competition assays showed only a small decrease in succinate formation, suggesting only very weak binding/inhibition by the D-K9me3 peptide (Fig. S29,30).

Our observations that KDMs do not (or extremely poorly) accept methylated D-lysine as substrates are consistent with their crystal structures complexed with histone peptides (Fig. 1C). Proximate positioning of the quaternary ammonium group of L-Kme3 to the active site iron is essential for efficient demethylation;[20] D-Kme3 presumably orients the N^{ε}-methyl group away from the iron, consequently leading to a non-productive binding mode. It is notable that human trimethyllysine hydroxylase, the first enzyme in the carnitine biosynthesis pathway, catalyses C-3 hydroxylation of the natural free L-Nε-trimethyllysine (obtained by the proteolytic degradation of L-N^{ε}-trimethyllysine containing histones), but not D-Nε-trimethyllysine.[21]

Overall, the results reveal that histone lysine methyltransferases, histone lysine demethylases and epigenetic readers efficiently modify or bind the (methyl) L-lysine, but manifest different levels of acceptance of (methylated) D-lysine as substrates/ligands. Our results imply that the protein-histone interactions, especially for the KDMs, critically determine the enzyme activity, whereas the associations between reader proteins and histones appear to be less sensitive to changes in lysine C_{α} stereochemistry. Notably, we did observe methyltransferase activity with some of the HKMTs with the D-configured substrates. Coupled with the lack of KDM activity and the observations of binding of D-residues by some reader domains, this raises the possibility that aberrant PTM patterns may occur should D-residues be produced in diseased or aged cells.

This work was supported by the ERC Starting Grant to J. M. (ChemEpigen-715691) and to A. K. (EPITOOLS-679479), the Netherlands Organization for Scientific Research (NCI-TA 731.015.202), Royal Society Dorothy Hodgkin Fellowship (A. K.), the Wellcome Trust, Cancer Research UK, the Engineering & Physical Sciences Research Council UK, the John Fell Oxford University Press (OUP) Research Fund, and the NIHR Oxford Biomedical Research Unit. We gratefully acknowledge Dr. D. McClymont for his support in developing the Matlab script.

Conflicts of interest

There are no conflicts to declare.

References

1 C. D. Allis, T. Jenuwein and D. Reinberg, Epigenetics, Cold Spring Harbor Laboratory Press, New York, 2007.

2 M. Tan, H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren and Y. Zhao, Cell, 2011, 146, 1016–1028.

3 L. Dai, C. Peng, E. Montellier, Z. Lu, Y. Chen, H. Ishii, A. Debernardi, T. Buchou, S. Rousseaux, F. Jin, B. R. Sabari, Z. Deng, C. D. Allis, B. Ren, S. Khochbin and Y. Zhao, Nat. Chem. Biol., 2014, 10, 365–370.

4 Z. Xie, J. Dai, L. Dai, M. Tan, Z. Cheng, Y. Wu, J. D. Boeke and Y. Zhao, Mol. Cell. Proteomics, 2012, 11, 100–107.

5 C. Qian and M.-M. Zhou, Cell. Mol. Life Sci., 2006, 63, 2755–2763.

6 S. M. Kooistra and K. Helin, Nat. Rev. Mol. Cell Biol., 2012, 13, 297–311.

7 S. D. Taverna, H. Li, A. J. Ruthenburg, C. D. Allis and D. J. Patel, Nat. Struct. Mol. Biol., 2007, 14, 1025–1040.

8 R. M. Hughes, K. R. Wiggins, S. Khorasanizadeh and M. L. Waters, Proc. Natl. Acad. Sci. U. S. A., 2007, 104, 11184–11188.

9 J. J. A. G. Kamps, J. Huang, J. Poater, C. Xu, B. J. G. E. Pieters, A. Dong, J. Min, W. Sherman, T. Beuming, F. M. Bickelhaupt, H. Li and J. Mecinović, Nat. Commun., 2015, 6, 8911.

10 L. S. Brunauer and S. Clarke, J. Biol. Chem., 1986, 261, 12538–12543.

11 R. J. Hopkinson, L. J. Walport, M. Münzel, N. R. Rose, T. J. Smart, A. Kawamura, T. D. W. Claridge and C. J. Schofield, Angew. Chem. Int. Ed. Engl., 2013, 52, 7709–7713.

12 N. Fujii, Biol. Pharm. Bull., 2005, 28, 1585–1589.

13 G. H. Fisher, EXS, 1998, 85, 109–118.

14 A. Poplawski, K. Hu, W. Lee, S. Natesan, D. Peng, S. Carlson, X. Shi, S. Balaz, J. L. Markley and K. C. Glass, J. Mol. Biol., 2014, 426, 1661–1676.

15 Y. Peng and E. Alexov, Proteins: Struct. Funct. Bioinf., 2016, 84, 1797–1809.

16 D. A. Case et al., AMBER 12, University of California, San Francisco, 2012.

17 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, J. Chem. Phys., 1983, 79, 926–935.

18 M. Schapira, Curr. Chem. Genomics, 2011, 5, 85–94.

19 R. J. Hopkinson, R. B. Hamed, N. R. Rose, T. D. W. Claridge and C. J. Schofield, ChemBioChem, 2010, 11, 506–510.

20 L. J. Walport, R. J. Hopkinson and C. J. Schofield, Curr. Opin. Chem. Biol., 2012, 16, 525–534.

21 A. H. K. Al Temimi, B. J. G. E. Pieters, Y. V. Reddy, P. B. White and J. Mecinović, Chem. Commun., 2016, 52, 12849–12852.

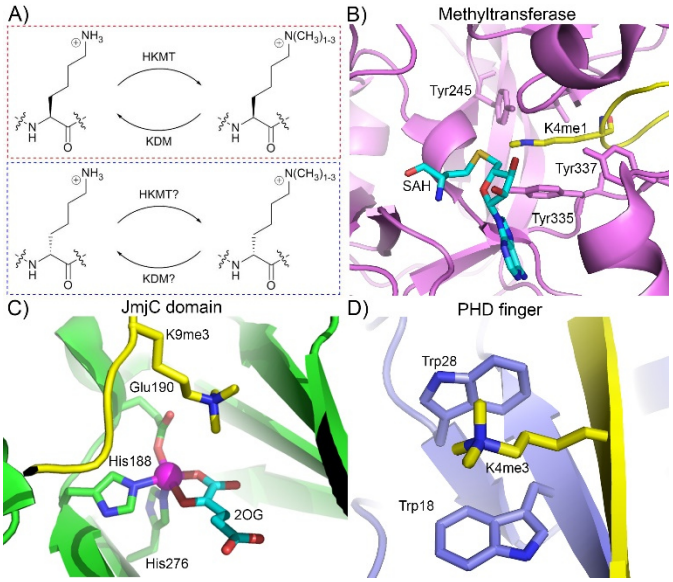

Fig. 1 A) N^{ε}-Methylation, demethylation and binding of histones. B) View from a SETD7 structure (magenta) complexed with H3K4me1 (yellow) and S-adenosylhomocysteine (cyan) (PDB: 1O9S). C) View from a $KDM4A_{JmjC}$ (green) structure complexed with H3K9me3 (yellow) and 2OG (cyan) (PDB: 2OQ6). D) View on $KDM5A_{PHD3}$ (blue) complexed with histone mimic peptide H3K4me3 (yellow) (PDB: 2KGI).

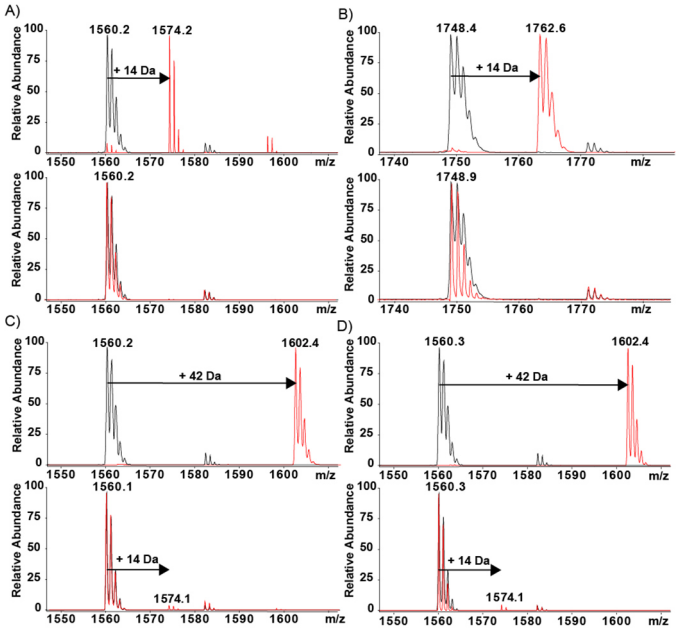

Fig. 2 MALDI-TOF MS assays for SAM-mediated methylation of L-lysine (top panel) and D-lysine (bottom panel) containing histone peptides by histone lysine methyltransferases. A) L-H3K4 and D-H3K4 with SETD7. B) L-H4K20 and D-H4K20 with SETD8. C) L-H3K9 and D-H3K9 with G9a. D) L-H3K9 and D-H3K9 with GLP. (black = starting peptide, red = after HKMT-catalysed reaction).

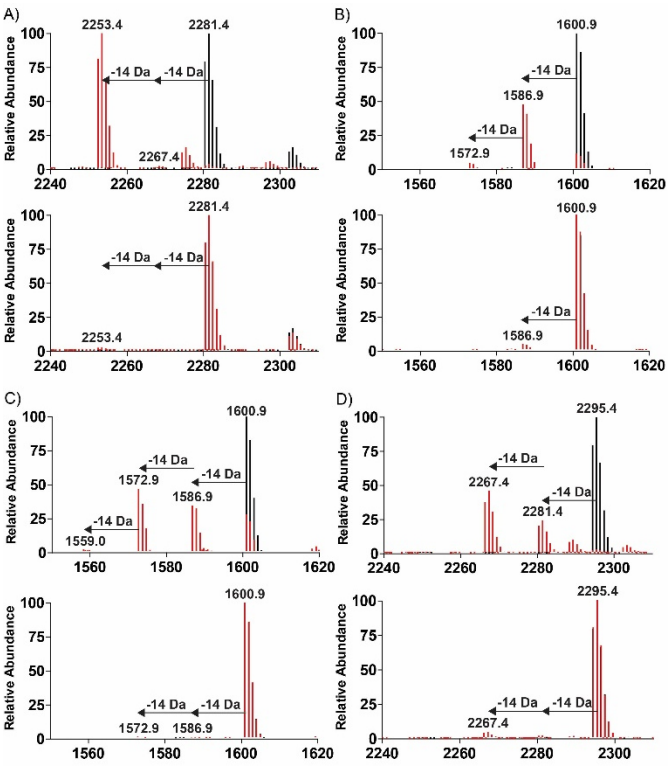

Fig. 3 Deconvoluted LC-MS data showing demethylation of methylated L-lysine (top panel) and methylated D-lysine (bottom panel) containing histone peptides in the presence of histone lysine demethylases, Fe(II), 2OG and ascorbate. A) L-H3K4me2 and D-H3K4me2 with KDM1A. B) L-H3K9me3 and D-H3K9me3 with $KDM4A_{JmjC}$. C) L-H3K9me3 and D-H3K9me3 with $KDM4D_{JmjC}$. D) L-H3K4me3 and D-H3K4me3 with $KDM5B_{JmjC}$. (black = starting peptide, red = after KDM-catalysed reaction). Conditions for JmjC-KDM assay: 10 μM 2OG, 10 μM Fe(II), 100 μM Asc.

Table 1 Thermodynamic parameters for binding of 10-mer histone peptides L-H3K4me3 and D-H3K4me3 (ART(L-Kme3/D-Kme3)QTARKS) to reader proteins. (Measured by ITC ± Standard Deviation (3–5 repeats).)

	L-H3K4me3				D-H3K4me3			
	K_d	$\Delta G°$	$\Delta H°$	$-T\Delta S°$	K_d	$\Delta G°$	$\Delta H°$	$-T\Delta S°$
KDM5A$_{PHD3}$	0.11	–9.5 ± 0.1	–11.1 ± 0.1	1 .6 ± 0.1	1.2	–8.1 ± 0.1	–7.8 ± 0.1	–0.3 ± 0.1
TAF3$_{PHD}$	0.082	–9.7 ± 0.1	–10.1 ± 0.1	0.4 ± 0.2	0.73	–8.4 ± 0.1	–8.0 ± 0.1	–0.4 ± 0.2
BPTF$_{PHD}$	0.44	–8.7 ± 0.1	–13.2 ± 0.1	4.5 ± 0.1	3.8	–7.4 ± 0.1	–7.7 ± 0.1	0.3 ± 0.1
SGF29$_{TTD}$	1.7	–7.9 ± 0.1	–7.7 ± 0.1	–0.2 ± 0.1	50	–5.9 ± 0.1	–2.3 ± 0.1	–3.6 ± 0.1
KDM4A$_{TTD}$	1.1	–8.1 ± 0.1	–13.1 ± 0.2	5.0 ± 0.2	39	–6.0 ± 0.2	–2.6 ± 0.1	–3.4 ± 0.2